Trade-Off Between Field-of-View and Resolution in the Thermal-Integrated Argus II System

Yingchen He¹, Natalie T. Huang², Avi Caspi³,⁴, Arup Roy⁴, and Sandra R. Montezuma²

¹ Department of Psychology, University of Minnesota, Minneapolis, MN, USA
² Department of Ophthalmology and Visual Neurosciences, University of Minnesota, Minneapolis, MN, USA
³ Department of Electrical and Electronic Engineering, Jerusalem College of Technology, Jerusalem, Israel
⁴ Second Sight Medical Products, Inc., Sylmar, CA, USA

Correspondence: Yingchen He, 75 E River Rd, Minneapolis, MN 55455, USA. e-mail: heccc340@umn.edu

Purpose: To investigate the effect of a wider field-of-view (FOV) of a retinal prosthesis on the users’ performance in locating objects.

Methods: One female and four male subjects who were blind due to end-stage retinitis pigmentosa and had been implanted with the Argus II retinal prosthesis participated (aged 63.4 ± 15.4). Thermal imaging was captured by an external sensor and converted to electrical stimulation to the retina. Subjects were asked to localize and to reach for heat-emitting objects using two different FOV mappings: a normal 1:1 mapping (no zoom) that provided 18° × 11° FOV and a 3:1 mapping (zoom out) that provided 49° × 35° FOV. Their accuracy and response time were recorded.

Results: Subjects were less accurate and took longer to complete the tasks with zoom out compared to no zoom. Localization accuracy decreased from 83% (95% confidence interval, 75%, 90%) with no zoom to 76% (67%, 83%) with zoom out (P = 0.07). Reaching accuracy differed between the two mappings only in one subject. Response time increased by 43% for the localization task (24%, 66%; P < 0.001) and by 20% for the reaching task (0%, 45%; P = 0.055).

Conclusions: Argus II wearers can efficiently find heat-emitting objects with the default 18° × 11° FOV of the current Argus II. For spatial localization, a higher spatial resolution may be preferred over a wider FOV.

Translational Relevance: Understanding the trade-off between FOV and spatial resolution in retinal prosthesis users can guide device optimization.

Introduction

The Argus II retinal implant system (Second Sight Medical Products, Inc., Sylmar, CA) is the only device that has obtained CE Mark and Food and Drug Administration approval for treating end-stage retinitis pigmentosa (RP).¹ Patients who are eligible for the device must meet the following criteria: (1) diagnosed with advanced RP, (2) have vision of bare-light or no light perception in both eyes, (3) had useful form of vision previously, (4) age 25 and older, and (5) are willing and able to undergo postsurgical visual rehabilitation therapy. There are more than 300 Argus II users worldwide.²

The prosthetic vision provided by Argus II is rudimentary and is below the level that can be measured with standard tests of visual function. To evaluate the visual function of these patients, special tasks were designed, ranging from localizing a high-contrast square on a computer screen³ to reading large letters or words,⁴ or more real-world related tasks such as sorting socks.⁵ Patients were able to achieve significantly better performance on various visual tasks with the aid of Argus II compared to using their residual vision,³⁶⁸ and their vision-related quality of life was also improved after implantation.⁹ The safety of the Argus II system has been demonstrated in a 5-year clinical trial.³

The current Argus II system translates the brightness level of the captured image into the amplitude of electrical stimulation on the retina. Without sufficient difference in brightness (i.e.,
contrast), the electrical stimulation does not induce a noticeable difference in visual perception and, thus, the wearer is unable to differentiate an object-of-interest against its background. Introducing the thermal camera as a complementary component to the regular visible light camera enables the wearers to differentiate heat-emitting objects from a cooler background. This feature would be helpful when the environment is cluttered or the difference in visible luminance is insufficient. Because heat-emitting objects can be relevant to entertainment and social interaction (e.g., humans and animals) or safety and mobility (e.g., avoiding hot stovetops and detecting the movement of cars), the thermal camera could be a helpful addition to the current Argus II system. Today, thermal imaging systems are used in many fields such as the military, medicine, and agriculture.

The effect of the integration of thermal imaging with Argus II has been reported by Dagnelie and colleagues (Dagnelie G, et al. IOVS. 2016;57:ARVO E-Abstract 5167; Dagnelie G, et al. IOVS. 2018;59:ARVO E-Abstract 4568). They tested the subjects’ ability to detect, localize, and count humans and hot objects in two conditions: (1) using the original Argus II system with its standard visible light camera and (2) using the Argus II system but substituting the visible light camera with a thermal camera. They found that the thermal-integrated Argus II system significantly improved performance compared to using the original Argus II system. Although the primary goal of the current study is not to evaluate the benefit of integrating the thermal sensor with the Argus II system, we used a similar thermal-integrated system (Fig. 1; described in Methods) as our research tool because it has the functionality for controlling the zoom of the camera input, which is not available in the Argus II visible light camera. We are interested in the following question: “How will an increased field of view (FOV) change the accuracy and response time of thermal target localization for Argus II wearers?”

We are interested in this question because the commercially available Argus II system provides an FOV of approximately 18° by 11° in front of the wearer or 20° diagonally, which is as large as the theoretical FOV subtended by the implant array on the retina. This FOV is limited and the patients need to use a head-scanning maneuver to view a larger area of interest. A restricted FOV impairs functional vision, especially mobility, and reduces quality of life, and a visual field of no greater than 20° is

Figure 1. Image of the thermal system with the Argus II system. OR, operating room; VPU, visual processing unit.
defined by the US Social Security Administration as legally blind. Therefore, it is natural to ask whether enlarging the FOV of the input video (thereafter referred to as the “input FOV”) can benefit Argus II users. In the past decades, various devices were designed for patients with tunnel vision to expand the FOV projected onto their retina. One strategy is to minify the scene, similar to zoom out using a camera. Given evidence that patients with tunnel vision can benefit from minifying telescopes, the Argus II wearers may also find a larger input FOV helpful. However, for the Argus II users, zooming out means that a larger visual angle is mapped onto the same implant region on the retina, effectively decreasing the spatial resolution of their prosthetic vision. This decrease in spatial resolution may offset any potential benefit brought by a larger input FOV.

To answer our research question, we used the thermal-integrated Argus II system to evaluate subjects’ performance to locate or reach for cups with hot water. Two different input FOV mappings were used: the natural no zoom mapping (1:1), where the FOV of the video input is mapped to correspond to the FOV subtended by the implant (18° by 11°); and a zoomed-out mapping (approx. 3:1), where the FOV of the video input was set to 49° by 35°. Subjects’ accuracy and response time were recorded and compared between the two input FOVs.

**Methods**

**Subjects**

Five subjects (one female and four male) who were implanted with the Argus II retinal prosthesis at the University of Minnesota were recruited in this study. The average age of the subjects (mean ± SD) was 63.4 ± 15.4 years old, ranging from 40 to 77. All subjects had bare light perception (BLP), with an average of 26.0 ± 9.0 years since their vision declined to BLP. The experience with the implant ranged from 3 months to 2.8 years, averaging 1.45 ± 1.25 years. Subjects signed informed consent to participate and a media consent form to allow publications of any photos or videos of them acquired during the experiment. The study was approved by the institutional review board at the University of Minnesota and conducted in accordance with the Declaration of Helsinki.

**Infrared Camera**

A Lepton forward looking infrared radiometer (FLIR) camera (FLIR Systems, Inc. Wilsonville, OR) with a resolution of 80 × 60 pixels and an FOV of 49° × 35° mounted in the center of a pair of dark-lensed eyeglasses was used to capture the thermal imagery (Fig. 1). The eyeglasses were connected to the thermal processor, which was then connected to the video processing unit (VPU) of the Argus II system.

The sensor output the temperature of the target in the instantaneous FOV of each pixel. The thermal processor converted the temperature of each pixel to a video (brightness) level based on a mapping curve with settable parameters (Fig. 2). The minimal and maximal temperatures used for the dynamic range of the thermal image, $T_{\text{min}}$ and $T_{\text{max}}$, were set using two parameters: $T_0$, which was the median temperature of the dynamic window; and $W$, which denoted the width of the temperature range. Objects with temperatures below $T_{\text{min}}$ were converted to 0 video level (black) and corresponded to no stimulation, whereas objects with temperatures above $T_{\text{max}}$ were converted to the maximum video level 255 (white) and corresponded to the highest-level stimulation. Objects with temperatures between $T_{\text{min}}$ and $T_{\text{max}}$ were converted to pixels whose brightness levels linearly correlated with their temperatures. During the experiment, the VPU was connected to the clinician programming system laptop for Argus II (Fig. 1) so that the video input from the camera and the stimulation patterns sent to the implant could be monitored by the experimenters, as shown in Figure 3B. Before the testing began, an experimenter adjusted the values of $T_0$ by using a slider on the thermal processor to best remove background noise based on the image on the laptop. $W$ was set to 2°C using a script in the thermal processor to provide high contrast between the object and the background.

The thermal processing unit sampled a region of interest within the FOV of the thermal imager and output a National Television System Committee...
Figure 3. (A) Demonstration of the two tasks. Top, localization task: a subject was detecting two cups of hot water using no zoom (left) versus zoom out (right). The cups could be presented either separately or together at one or two of the three locations: right, center, or left, respectively. Bottom, reaching task: the subject was reaching to a cup of hot water using no zoom (left) versus zoom out (right). The location of the cup could be right, left, or center relative to the subject. (B) Thermal camera video output and prosthesis stimulation patterns (screenshots from the laptop). Top, two cups of hot water together. Bottom, one cup of hot water. Left, no zoom. Right, zoom out. Within each panel, the grayscale image on the left shows the video input from the thermal processor, where brighter pixels correspond to a higher temperature. The image with individual dots on the right side of each panel represents the electrical stimulation patterns delivered by the implant, where each dot corresponds to one electrode and brighter color means stronger stimulation with respect to that electrode’s threshold.
Signal that contained the sampled region of interest. The region of interest was sent to the VPU using the splice cable. Two different regions of interest were tested: a normal 1:1 mapping (no zoom), which sampled $18^\circ \times 11^\circ$ of FOV from the center of the camera input; and a 3:1 mapping (zoom out), which surveyed the entire $49^\circ \times 35^\circ$ of FOV of the thermal camera. The location and the orientation of the region of interest within the FOV of the imager was set using a look-up table that was saved on the microSD card inside the thermal processor. The look-up table was generated by the experimenter using a MATLAB script.

**Protocol**

The participants sat in front of a table and were asked to perform two functional tasks. In the Localization task, subjects were asked to identify the locations of two cups of hot water, which could be located either together or separately at one or two of the three locations: right, center, or left (Fig. 3A, upper panels). In the reaching task, subjects were asked to reach for a cup of hot water that was located at one of the three locations: right, center, and left (Fig. 3A, lower panels). Successful reaching requires the subject to correctly understand the spatial correspondence between the visual information and the real-world locations. The reaching task tested how performance was influenced when the zooming condition changed this spatial correspondence. To minimize the risk of injury from the cup of hot water, the subjects were instructed to reach out slowly so as not to knock the cup over. The experimenter closely monitored the subjects in order to intervene in case of any potential danger. No intervention was needed and none of the subjects reported any discomfort caused by touching the cup.

Figure 3B shows the thermal camera output and prosthesis simulation patterns. Although there was little contrast between the cups and the tablecloth in visible luminance level (Fig. 3A), the cups were clearly visible in the thermal imagery (Fig. 3B), demonstrating one case where the thermal camera can be helpful. The 1:1 and 3:1 mappings were used for both tasks. The testing order of the two FOV mappings was balanced across tasks and subjects (e.g., S1: localization—no zoom, localization—zoom out, reaching—zoom out, reaching—no zoom; S2: localization—zoom out, localization—no zoom, reaching—no zoom, reaching—zoom out). Before each task block, subjects went through a short training to use the system to scan for their own hands and to complete a few practice trials of the task. Each task was repeated for 20 trials for each mapping, except for our first subject S1. The protocol for S1 only included 10 trials for task 2; the number of trials was changed to 20 for subsequent subjects to increase statistical power. The set of cup configurations was the same across subjects and across zoom conditions, with randomized trial sequences. The accuracy and response time were recorded and compared for the two different FOV mappings.

**Statistical Analysis**

The number of correctly located cups in the localization task (coded as 0/1/2) and the success/failure to reach a cup in the reaching task (coded as 0/1) were fitted with cumulative link mixed models. Response time data were fitted with linear mixed-effect models. FOV mapping (no zoom/zoom out) was modeled as the fixed effect, representing the main effect of zoom. Subject and cup configuration were modeled as random effects, accounting for the differences between individual subjects and trials. When computing average accuracy, the number of correctly located cups in each trial was divided by the total number of cups (2 for localization and 1 for reaching). Response time was log-transformed before statistical testing to follow a normal distribution. When reporting results, we transformed the response time back to its original unit (seconds) for easier understanding. All tests and figures were produced using R. Significance level was defined as $P < 0.05$.

**Results**

Subjects’ accuracy and response time to complete the two tasks, with the two different mappings, are presented in Tables 1 and 2 and Figure 4. Below, we discuss the effect of zoom on accuracy and response time.

**Accuracy**

Table 1 and Figure 3A show the subjects’ average accuracy for the two tasks with no zoom and zoom out. Subjects were able to perform both tasks above chance level by using the thermal camera regardless of the zoom. In the localization task, four out of the five subjects had higher average accuracy to identify the locations of two cups when using no zoom compared to zoom out. On average, the accuracy decreased from 0.83 (95% confidence interval, 0.75, 0.90) for no zoom to 0.76 (0.67, 0.83) for zoom out ($P = 0.07$).
In the reaching task, four of the five subjects were able to complete the reaching task at 100% accuracy in both zooming conditions. Only one subject, S4, showed higher accuracy using no zoom compared to zoom out. Mixed-effect modeling also revealed a significant effect of zoom ($P = 0.0354$). However, the magnitude of the fixed effect was trivial compared to the random effect estimated by the model. That is, the variation between subjects and between different cup configurations was much larger compared to the difference caused by zooming. Thus, the effect of zoom could not be determined in our study due to the ceiling effect (100% accuracy in 4/5 subjects). The one subject with higher accuracy using no zoom (S4) had poorer overall performance in both the localization and reaching tasks, and this could explain why he was less affected by the ceiling effect. As an observation, the localization task was more challenging for all subjects compared to the reaching task regardless of the zoom used. This was likely because, for the first task, two objects needed to be localized and they could appear separated (requiring greater scanning) or in the input FOV simultaneously, causing confusion. In contrast, for the second task, the subject was only required to localize one object, making it easy to interpret the mapping between visual input and spatial location. Future studies could use more detailed quantification of the reaching behavior to examine the effect of zoom, such as the reaching speed or the maximum deviation or use a more difficult reaching task altogether in order to reduce the ceiling effect.

### Response Time

Table 2 and Figure 3B show the subjects’ response time with the two mappings. Overall, most subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>No Zoom</th>
<th>Zoom Out</th>
<th>Accuracy Changea</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.90 (0.68, 0.97)</td>
<td>0.70 (0.47, 0.86)</td>
<td>−0.20 (−0.49, 0.09)</td>
</tr>
<tr>
<td>S2</td>
<td>0.85 (0.62, 0.95)</td>
<td>0.80 (0.57, 0.92)</td>
<td>−0.05 (−0.33, 0.23)</td>
</tr>
<tr>
<td>S3</td>
<td>0.88 (0.65, 0.96)</td>
<td>0.85 (0.62, 0.95)</td>
<td>−0.03 (−0.26, 0.21)</td>
</tr>
<tr>
<td>S4</td>
<td>0.80 (0.57, 0.92)</td>
<td>0.53 (0.31, 0.73)</td>
<td>−0.28 (−0.61, 0.06)</td>
</tr>
<tr>
<td>S5</td>
<td>0.75 (0.52, 0.89)</td>
<td>0.93 (0.70, 0.98)</td>
<td>0.17 (−0.10, 0.45)</td>
</tr>
<tr>
<td>Average</td>
<td>0.83 (0.75, 0.90)</td>
<td>0.76 (0.67, 0.83)</td>
<td>−0.07 (−0.20, 0.05)</td>
</tr>
</tbody>
</table>

Values for S1–S5 indicate average accuracy of all trials for each subject, rounded to two decimal places. Parentheses indicate 95% confidence intervals, capped at 1.

a Accuracy change, zoom out – no zoom. Asterisks indicate significant changes. *0.01 ≤ P < 0.05.

In the reaching task, four of the five subjects were able to complete the reaching task at 100% accuracy in both zooming conditions. Only one subject, S4, showed higher accuracy using no zoom compared to zoom out. Mixed-effect modeling also revealed a significant effect of zoom ($P = 0.0354$). However, the magnitude of the fixed effect was trivial compared to the random effect estimated by the model. That is, the variation between subjects and between different cup configurations was much larger compared to the difference caused by zooming. Thus, the effect of zoom could not be determined in our study due to the ceiling effect (100% accuracy in 4/5 subjects). The one subject with higher accuracy using no zoom (S4) had poorer overall performance in both the localization and reaching tasks, and this could explain why he was less affected by the ceiling effect. As an observation, the localization task was more challenging for all subjects compared to the reaching task regardless of the zoom used. This was likely because, for the first task, two objects needed to be localized and they could appear separated (requiring greater scanning) or in the input FOV simultaneously, causing confusion. In contrast, for the second task, the subject was only required to localize one object, making it easy to interpret the mapping between visual input and spatial location. Future studies could use more detailed quantification of the reaching behavior to examine the effect of zoom, such as the reaching speed or the maximum deviation or use a more difficult reaching task altogether in order to reduce the ceiling effect.

### Response Time

Table 2 and Figure 3B show the subjects’ response time with the two mappings. Overall, most subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>No Zoom</th>
<th>Zoom Out</th>
<th>Percent Changea</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>25 (20, 31)</td>
<td>29 (24, 34)</td>
<td>15 (−9, 46)</td>
</tr>
<tr>
<td>S2</td>
<td>13 (11, 15)</td>
<td>38 (34, 41)</td>
<td>197 (148, 256)</td>
</tr>
<tr>
<td>S3</td>
<td>14 (11, 17)</td>
<td>24 (21, 27)</td>
<td>75 (44, 112)</td>
</tr>
<tr>
<td>S4</td>
<td>38 (28, 50)</td>
<td>47 (41, 54)</td>
<td>25 (−6, 67)</td>
</tr>
<tr>
<td>S5</td>
<td>31 (25, 38)</td>
<td>25 (21, 29)</td>
<td>−20 (−38, 4)</td>
</tr>
<tr>
<td>Average</td>
<td>22 (14, 35)</td>
<td>31 (24, 41)</td>
<td>43 (24, 66)***</td>
</tr>
</tbody>
</table>

Values for S1–S5 indicate average response time in seconds of all trials for each subject, rounded to whole seconds. Parentheses indicate 95% confidence intervals.

a Percent change for response time is (zoom out – no zoom)/no zoom. Asterisks indicate significant changes. ***P < 0.001.
spent a longer time to locate or to reach for the target using zoom out, with some variation across individuals. For the localization task, four of the five subjects had longer average response time for zoom out compared to no zoom (Table 2). The average time for subjects to identify the locations of two cups of hot water was 22 seconds for no zoom (95% CI, 14, 35). For zoom out, the time increased to 31 (24, 41) seconds. The linear mixed-effect model revealed a significant effect of zoom ($P < 0.001$). According to the model, subjects would spend 43% (24%, 66%) more time on the task with zoom out.

In the Reaching task, three of the five subjects took longer to complete the task with zoom out (Table 2). The average time to reach for the cup was 10 seconds (95% CI, 7, 15) with no zoom and 12 (8, 18) seconds with zoom out ($P = 0.055$), and response time was estimated to increase by 20% (0%, 45%) from no zoom to zoom out.

**Discussion**

Previous studies have shown that the Argus II retinal prosthesis can create some artificial vision for the wearers and improve visually guided fine hand movements. In the current study, we use a thermal-integrated Argus II system to evaluate the trade-off between the FOV and the spatial resolution on the wearer’s ability to correctly localize and to reach for targets.

A previous study suggested that object recognition requires a larger FOV when spatial resolution is reduced. Our initial hypothesis was that a wider input FOV would allow Argus II wearers to detect...
the targets faster, thus decreasing the response time. However, successful target localization depends on effective stimulation that exceeds the subject’s perceptual threshold; in other words, the phosphenes elicited by the electrical stimulation must be bright and big enough to be perceived by the subject. When a wider input FOV was mapped onto the area of the retinal implant, the spatial resolution of the stimulation decreased, so the same object decreased in size in the video input and, therefore, activated fewer electrodes. The perceptual threshold of each electrode may depend on factors such as electrode-retina distance and ganglion cell density, and, therefore, the sensitivity to electrical stimulation can vary between electrode loci. As a result, if the target activated fewer electrodes, there would be less chance for the stimulation to be perceived. In addition, because the video input refreshes at 30 Hz, there may be a delay of approximately 33 ms from video capture to the electrical stimulation. If the scanning motion was too fast, localization error could occur due to this delay. With the wider input FOV, subjects often missed the target during their initial head scanning and had to scan multiple times before successfully locating the targets. We did notice that one subject, S5, showed a different pattern from other subjects (Fig. 4, purple dotted line in panels A2 and B2). S5 reacted faster and was more accurate when using zoom out compared to no zoom. We noted that S5’s retina was more sensitive to electrode stimulation according to the most recent implant fitting measurement. It is possible that when more electrodes are effectively stimulated, the adverse effect of reduced spatial resolution can be alleviated and the subject may benefit more from a wider input FOV. Similarly, when prescribing minifying telescopes to RP patients, good visual acuity (20/30 or better) was recommended in order to achieve the maximum efficacy. As experienced by low vision care providers, field expanders have not been widely adopted by RP patients despite positive findings in several small studies, mostly due to the trade-off between visual acuity and field size (personal communication, 2019).

One more factor that could influence task performance was experience or practice. The spatial correspondence provided by the wider input FOV between retinal stimulation and the world was new to the subjects. The subjects have substantial experience with the normal input FOV but only had brief training and familiarization with the wider input FOV. It could, thus, be argued that the lack of experience may explain the poor performance with the wider input FOV. However, an earlier study showed that when an eye tracker was integrated with the Argus II system to shift the input FOV corresponding to eye movements, the patients showed instantaneous improvements in localizing a target. This finding supported that the patients are able to learn a new method of manipulating the input FOV quickly and benefit from it. In addition, a longitudinal study showed that the Argus II wearers’ performance on a target localization task did not improve over the course of 5 years after the device was activated (inferred from Fig. 2A in Reference 3), suggesting that additional practice does not necessarily lead to improved performance. Without further experimental testing, it is unclear whether more practice with the wider input FOV would change our negative finding.

Overall, our results showed that Argus II wearers were able to use the information provided by the thermal camera and could find heat-emitting objects using the current 11° by 18° input FOV with greater accuracy and speed than when using a wider input FOV. The results are applicable to the thermal camera, but it remains to be tested whether the conclusions could generalize to the visible light camera where the video input contains a high-contrast, uncluttered target (for example, using the Argus II camera to locate a dark door against a white wall or a white plate on a black tablecloth). For object localization, a higher spatial resolution may be preferred over a wider input FOV.

Acknowledgments

The authors thank Jessy Dorn for her generous help in critically reviewing the manuscript. Sponsorship and article processing charges for this study were funded by the Center for Applied Translational Sensory Science at the University of Minnesota, the Minnesota Lions Vision Foundation, the VitreoRetinal Surgery Foundation, and an unrestricted grant to the Department of Ophthalmology and Visual Neurosciences from the Research to Prevent Blindness (RPB), New York, NY, USA.

Disclosure: Y. He, None; N.T. Huang, None; A. Caspi, Second Sight Medical Products, Inc. (C,P); A. Roy, Second Sight Medical Products, Inc. (I,E,P); S.R. Montezuma, None
References


